人工智能方法在探究小学生作业作弊行为及其关键预测因子中的应用*

赵立[1#] 郑怡[2#] 赵均榜[3] 张芮[1,4] 方方[5] 傅根跃[1] 李康[6]

([1]杭州师范大学心理学系, 杭州 311121)

([2]杭州师范大学经亨颐教育学院, 杭州 311121)

([3]浙江师范大学儿童发展与教育学院, 杭州 311231)

([4]杭州市夏衍小学, 杭州 311121)

([5]北京大学心理与认知科学学院, 行为与心理健康北京市重点实验室, 北京 100871)

([6]加拿大多伦多大学, 安大略教育研究所, 安大略 M5R 2X2)

摘 要 小学生作业作弊是心理学领域忽略已久的研究重点，机器学习是数智时代新兴的人工智能科学。笔者对 2,098 名 2 至 6 年级小学生进行问卷调查，采用机器学习法，考察个体认知、道德判断、同伴行为，及性别、年级、成绩等因素对小学生作业作弊行为的影响。结果表明：集成机器学习模型对小学生作业作弊预测准确率(AUC 均值)达 80.46%；对作业作弊预测效应最强的四个因素依次为个体对作业作弊的接受程度、观察到同伴作弊的普遍性和频率，及其自身成绩。

关键词 小学生，诚信，学业作弊，作业作弊，机器学习，预测，同伴行为

收稿日期: 2022–07–14

* 国家自然科学基金项目(32171060)和浙江省哲学社会科学规划领军人才培育专项课题(21QNYC18ZD)资助。

赵立和郑怡为共同第一作者

通讯作者：赵立, E–mail: zhaoli@hznu.edu.cn; 李康, E-mail: kang.lee@utoronto.ca

1 引言

学业作弊是古往今来困扰教育工作者的难题。它是指学生为获得理想的学业成绩或排名而采取的秘密且有意破坏学业诚信准则的行为，其重要表现形式之一为作业作弊。虽然自“双减”以来，我国小学生的课业负担明显减轻，但不可否认，较之国外，“作业”依然是我国义务教育阶段小学生最主要的学业任务，作业作弊行为在小学生中仍普遍存在。小学阶段是个体道德意识形成及道德行为习惯塑造的关键期，这一阶段学生做出的作业作弊行为既会导致其学业成绩不佳进而在考试中作弊，久而久之，又容易使其作弊行为习惯化。并且，对儿童期的作弊行为若不重视并加以及时引导和干预，很容易导致个体在成人后表现出更严重的问题行为，如违法乱纪或犯罪(Williams & Williams, 2012)。且“双减”政策在强调“减负”的同时，也明确了提高作业质量的需求，并提倡杜绝“唯成绩论”。可以预期，相较于考试成绩，小学生的作业质量等平时学业表现在未来的学业评价体系中可能更为重要。为此，本研究就作业作弊行为在小学阶段的发展及其关键性影响因素进行深入考察，以期为探索作弊的发展机制以及提出有针对性的早期干预方法提供科学依据。

在学术界，有关学业作弊的研究已有近百年历史(Hartshorne & May, 1928)。纵观现有的实证研究，绝大多数都来自国外，且研究对象几乎全部聚焦于大学生和中学生群体(参见 Anderman & Midgley, 2004; Cizek, 1999; Hrabak et al., 2004)，极少有研究关注小学生群体(Hartshorne & May, 1928)。虽然，早在 1928 年，Hartshorne 和 May 的研究就已揭示小学阶段儿童会自发地做出学业作弊行为。但在此之后，因西方教育体制改革导致国外小学阶段的考试和作业一度被取消，故针对小学生作弊的实证研究也一度停滞。与此同时，在研究内容方面，国外学业作弊的研究主要集中于考试作弊(Bong, 2008; Freire, 2014)，针对作业作弊的研究非常少。可以说，到目前为止尚未有研究专门考察过小学生的作业作弊行为。

与国外相比，国内有关学业作弊的实证研究起步更晚，且研究对象全部集中于大学生

群体(如舒首立等, 2018; 易晓明, 2021)，针对小学生作弊(尤其是作业作弊)的实证研究尚属空白。虽然有关大学生学业作弊的研究结果可以为小学生作业作弊研究提供一定的思路和借鉴。但与大学生相比，小学生的道德认知发展水平更低，故针对前者的研究结果是否适用于小学生群体尚不明确。鉴于此，本研究将立足我国国情，以国内外现有实证研究结果为基础，重点考察影响当前我国小学生作业作弊行为的可能因素，并从中寻找关键性影响因素。本研究结果旨在为建构儿童诚信行为发展的理论模型，以及道德行为发展理论的完善和本土化提供重要的科学依据。

具体说来，本研究将以 2 至 6 年级小学生为研究对象，采用大样本问卷调查法和机器学习(Machine Learning, Pedregosa et al., 2011)相结合的方法，考察小学生作业作弊行为的年级发展趋势，重点分析不同因素对作业作弊的影响及其相对重要性。

机器学习是数智时代新兴的数据分析方法，也是人工智能的核心方法，它依赖于计算机算法，旨在通过数据分析和建模实现对人类行为的模拟。近年来，不少心理学研究者已将机器学习应用于诸如情绪(Just et al., 2017)和精神病理学等方面的研究(如 Bartlett et al., 2014; Livieris et al., 2018)，也有少数研究者开始将该方法用于儿童研究(Bruer et al., 2019; Zanette et al., 2016)。与传统的统计建模相比，机器学习具有以下四个方面的优势。

首先，机器学习对原始数据的处理有助于提高最终模型的外部效度。传统分析方法(如，广义线性模型 GLM、广义估计方程 GEE，等等)通常将所有数据统一纳入分析，往往导致模型过度拟合，所获模型的概化能力即外部效度较差。与之不同，机器学习(往往需要较大的样本量)事先将数据分为三个子集：训练集(train set)、测试集(test set)和留出集(holdout set)。首先用训练集数据来拟合模型，然后用测试集数据来检验模型。再对这两个子集进行重组，重新随机划分为新的训练集和新的测试集，进行模型拟合和验证。这个过程将重复多次，并产生多个可预测结果变量的模型。最后，用事先预留的留出集数据对这

些模型的预测力做进一步评估，即验证这些模型的预测力能否概化到新的数据中。这种验证实现了对模型外部效度的评估(Campbell, 1986)。

其次，机器学习囊括了多种算法，可被灵活应用于处理各类复杂多变的变量间关系。本研究将采用近年来广受关注的集成学习法(Ensemble Learning, 参见 Ykhlef & Bouchaffra, 2017)。该方法首先通过不同机器学习算法对同一样本群体进行训练。具体算法包括描述变量间线性关系的逻辑回归(Logistic Regression, Yarkoni & Westfall, 2017)，描述变量间非线性关系的多层感知机(Multilayer Perceptron, MLP)、极端梯度提升(eXtreme Gradient Boosting, XGBoost)和随机森林(Random Forest)等(参见, Golino et al., 2014)。随后，将所有算法的训练结果整合起来。集成法的优势在于能够集各种算法之长，最大程度上对预测变量和结果变量的关系进行解释，从而使最终模型的预测效果达到最佳。

第三，机器学习可通过 Shapley 值，量化不同影响因素在整个模型中的相对重要性。Shapley 值于 1953 年由诺贝尔经济学奖获得者 Lloyd Shapley 提出(Shapley, 1953)，主要用于衡量各个预测变量对结果变量的相对贡献大小，是机器学习结果的重要参考指标之一(Smith & Alvarez, 2021)。该值越大，代表该预测因素对结果变量的预测力越强；反之，则代表该因素的预测力越弱。若该值趋近于零时，则代表该因素对结果变量几乎不存在有效预测力。

最后，机器学习能够将所有预测变量对结果变量的预测效应进行量化，从而创建一个关于结果变量(如“小学生作业作弊的概率”)的预测模型。模型建立后，若将某个学生在相应预测变量上的得分输入模型，即可输出该学生作业作弊的可能性(即作弊的概率)。依据这一结果，能够初步预测学生有多大的可能性会在作业中作弊，以便于教师或家长进行必要的关注，并结合学生的实际表现采取针对性的教育和干预措施。

基于上述机器学习的优势，本研究将重点考察部分因素(即预测因素)对小学生作业作

弊行为的影响。在影响因素选取方面，由于国内外对小学生作业作弊行为影响因素的研究几乎为空白，故本研究主要参考了 Murdock 和 Anderman(2006)的学业作弊动机模型，以及以往针对大中学生作业作弊行为的研究成果。

Murdock 和 Anderman(2006)的学业作弊动机模型是目前为止在作弊(尤其是考试作弊)研究领域影响力最大且应用最广的模型之一 。该模型整合了一系列以大中学生为主要研究对象的研究成果，从“作弊的代价”、“作弊的目的”及“作弊前对自我和结果的认知(我能做到吗)”三个方面对学业作弊动机进行了划分。其中，“作弊的代价”主要包括作弊后被抓的后果、个体道德水平、同伴作弊的后果、关于诚信的规章制度等等；“作弊的目的”包括同伴压力、智力、课堂氛围等因素；“作弊前对自我和结果的认知”则主要为自我效能、及个体对结果的预期(Murdock & Anderman, 2006)。

依据上述学业作弊动机模型，在综合考虑了研究成果未来在小学中的可实践性和可推广性，同时参考了以往有关大中学生学业作弊的实证研究结果，本研究对可能影响小学生作业作弊行为的因素(即小学生作业作弊的预测因素)进行了筛选，重点包括以下几方面。

第一，作弊后果的严重性。一方面，在 Murdock 和 Anderman (2006)的学业作弊动机模型中，作弊的后果被看作是学业作弊的主要动机之一。随后更多的实证研究也表明，作弊的后果严重性对大中学生学业作弊行为存在影响(Kam et al., 2017; McCabe & Treviño, 1997; Molnar & Kletke, 2012)。另一方面是基于实践证据。众所周知，“惩罚”一直被我国教育实践工作者视作防止作弊发生的“最有效/最佳措施”。然而，无论是古时将科举舞弊者量刑入罪，还是如今教育主管部门及各级各类学校制定的针对作弊的处罚制度(如取消考试成绩、取消考试资格、通报批评等)，均只针对考试作弊。相较于考试作弊，目前少有学校对学生的作业作弊行为予以重视和深究。这既是因为作业作弊相对不容易被发现，又是由于其频繁性和参与人数可能更高。正是因为作业作弊往往被看成是一种风险极低且无不良后

果的违规行为，导致了作业作弊的普遍性和频率往往较高。因此，本研究假设，作弊后果严重性程度是影响小学生作业作弊行为的首要因素。

第二，作弊的可接受性。Murdock 和 Anderman(2006)的学业作弊动机模型认为，个体的道德水平是影响学业作弊的另一个重要因素。以往针对大学生和中学生的实证研究表明，个体的道德意识水平影响其对作弊可接受程度的判断(Cheung et al., 2016; Lee et al., 2020)，以及他人对作弊接受程度的判断(Ives & Giukin, 2020)。错误的判断很可能导致个体美化或弱化学业作弊行为。故本研究假设，个体自身以及他人对作弊的可接受性是影响小学生作业作弊行为的重要因素。

第三，小学生对作业作弊预防措施的有效性评价。Murdock 和 Anderman(2006)的学业作弊动机模型指出，学校与诚信相关的规章制度对作弊行为具有一定的影响。这一观点在之后的实证研究中也得到了证实(Ramberg & Modin, 2019)。但另一些调查发现，学生对于学校制定的学业诚信政策或许并不十分了解(Bretag et al., 2014; Gullifer & Tyson, 2014)，这可能是导致学业作弊屡禁不止的潜在原因之一，即相关政策并未发挥应有的约束作用。因此，本研究假设，学生对作弊预防措施的评价，能够有效预测其作业作弊行为。

第四，同伴作弊行为。以往在大中学生中开展的实证研究曾表明，观察到同伴做出作弊行为，可能会导致个体对作弊这一行为进行合理化(Jurdi et al., 2011; McCabe & Abdallah, 2008)，进而影响其自身参与作业作弊行为的概率(Hrabak et al., 2004)。而另一方面，观察到同伴参与作弊行为，也可能使观察者产生“不公平感”，进而更倾向于作弊。故本研究假设，同伴作弊行为能够有效预测小学生的作业作弊行为。

除了上述预测变量外，本研究还考察了包括年龄、性别、以及自我报告的自身学业成绩水平等一系列人口统计学或个人背景变量对小学生作业作弊行为的预测作用。其中，以往针对大中学生的研究发现，成绩较差的学生更易为了获得好成绩而参与学业作弊行为

(Newstead et al., 1996; Özcan et al., 2019)，且自我效能感低(即对自身能力认知较差的人)的学生作弊的可能性也越高(see Murdock & Anderman, 2006)；而诸如年龄、性别等人口统计学变量(Błachnio, 2019; Cizek, 1999; Freire, 2014; Jurdi et al., 2011)与个体的学业作弊行为也存在相关。故本研究假设，上述变量对小学生作业作弊行为存在显著预测作用。

综上，本研究将采用问卷调查法，结合机器学习技术，考察(1)作业作弊对个体可能造成的后果及其严重性的感知、(2)个体自身及感知到同伴对作业作弊的接受程度、(3)对各种预防作业作弊措施的有效性评价、(4)感知到同伴的作业作弊行为，以及(5)年龄、性别等人口统计学变量对小学生作业作弊行为的影响；在此基础上，构建机器学习模型，采用集成法分析比较上述因素对作业作弊行为的预测效应。

2 方法

2.1 被试

本研究事先经过了杭州师范大学学术伦理委员会的伦理审查，并征得了学生所在学校及学生家长的知情同意。在浙江省某地级市选取了三所不同类型的小学：普通公办小学(以下简称“公办小学”)、含外来务工人员较多的公办小学(以下简称“民工子弟小学”)及民办小学各 1 所，选择 2 至 6 年级，从每个年级中随机抽取若干班级，进行整班问卷调查(考虑到 1 年级小学生的识字水平和阅读理解能力均十分有限, 且其家庭作业很少, 故本次调查未将其纳入; Zhang, 2019)。共计 2300 名小学生参与了本次问卷调查。其中，157 份问卷完成率不足 70%，45 份问卷结果变量相关题目未作答，予以剔除。最终得到有效问卷 2098 份。有效样本的平均年龄为 10.04±1.40 岁(男生占 53%)，全体学生均为汉族，年级、性别及学校等分布情况见下表。

表 1 2098 名有效被试的部分人口统计学信息情况

年级	年龄（岁）		性别		学校			合计
	均值	标准差	男	女	普通公办	民工子弟小学	民办小学	（人）

2 年级	7.86	0.60	205	192	210	83	104	397
3 年级	8.79	0.40	209	189	200	84	114	398
4 年级	9.79	0.33	209	176	186	85	114	385
5 年级	10.74	0.33	240	211	183	131	137	451
6 年级	11.75	0.34	253	214	155	173	139	467
							合计(人)	2098

2.2 研究工具

2.2.1 问卷编制与施测

考虑到国内外尚未有专门针对小学生作业作弊行为的问卷，故本研究采用自编的问卷进行测量，问卷在以往针对大中学生学业作弊的相关研究的基础上编制而成。问卷编制过程分为三个阶段。

第一阶段是访谈。基于 Lim 和 See(2001)对大学生学业作弊的现状、态度、后果严重性等的问卷调查结果，编制访谈提纲。访谈主要针对学生和教师两类群体开展。在学生层面，访谈内容主要涉及当前小学生作弊(尤其是作业作弊)行为的现状、表现形式及其对作弊的认知和态度等；在教师层面，除上述内容外，访谈还涉及学校、家庭在应对学生作弊问题方面的一些方法和措施，及其所取得的成效。来自上述三所小学中的 39 名学生(这些受访学生不再参与后续的正式问卷调查)和 9 名教师参与访谈。访谈全程录音，并由两位不知研究目的的心理学硕士生进行转录和编码。访谈结果表明：作业作弊在 2 年级小学生中就已存在，主要表现为抄标准答案和抄同学作业等；小学生普遍认为的作弊的严重后果主要涉及被老师或家长批评，被同学嘲笑等；其认为能够减少作弊的有效措施主要有努力学习掌握知识点、加大惩罚力度和告知家长等。

第二阶段为小范围试测。在访谈的基础上，进一步参考 Bucciol 等人(2017)有关大学生学业作弊的研究问卷，编制本研究的初测问卷。在上述三所学校，随机选取来自 2 年级、4

年级和 5 年级的共 158 名学生(初测被试同样不参与后续的正式问卷调查)进行小范围试测。根据问卷作答情况，结合试测过程中学生的现场反馈，对初测问卷进行修订，并对部分表述不清，难以理解，或有疑问的条目及选项进行了修订和完善。邀请两位心理测量学方面的专家对修改后的问卷进行了审定，最终形成本研究的正式问卷。

第三阶段为正式调查。在上述三所学校，从 2 至 6 年级的各个班级中按一定比例抽取学生进行问卷调查(为保证被试代表性，抽样时确保参与调查学生的学业成绩分布较为广泛)。问卷调查采用一对多的方式，在各个班级内独立进行。问卷现场发放，当场收回。为最大程度地消除学生的顾虑，问卷采用匿名的形式，且问卷发放、填写和回收的整个过程均由研究助理出面组织，班主任或任课教师均未在现场，也未参与整个问卷调查过程。

2.3 正式问卷构成

正式问卷主要对两部分内容进行测量。一是研究的关键结果变量——作业作弊行为，二是预测变量，即作业作弊行为影响因素，包括个体心理变量和人口统计学变量两个方面。

2.3.1 结果变量——作业作弊行为(简称“作弊行为”)

请作答者就自身“在平时学习中做过类似抄别人的作业/抄答案/让别人帮忙写作业等的行为吗？”，从“1”(从不)到“5”(非常多)进行 5 点利克特量表评分。

笔者在后续的数据分析中发现，全体样本对该问题的评分呈明显的正偏态(即评分为“1”的学生比例很高，评分为“2”～“5”的学生比例相当且均较低)，无法按照近似正态分布处理。为避免统计偏差，遂将结果变量重新编码为二分变量，其中将评分为“1”者编码为“无作弊行为者”，以 0 计；将评分为“2”～“5”者编码为“有过作弊行为者”，以 1 计。

2.3.2 **预测变量(影响因素)的测量**

(1)**对作业作弊行为可能导致后果的严重性评价(简称“后果严重性”)**

根据访谈结果，列举了 5 种作业作弊行为(Cronbach's α 系数为 0.787)可能导致的后果(如“老师批评”等)，请作答者就每一后果的严重性程度，从“1”(一点也不严重)到“5”(非常严重)进行 5 点利克特量表评分。

(2)**对作业作弊行为的可接受性(简称“可接受性”)**

包括两个条目：a. 小学生自身对作业作弊行为的可接受程度(简称“*自我可接受性*”)；b.感知到同伴对作业作弊行为的可接受程度(简称“*同伴可接受性*”)。请作答者从“1”(完全不能接受)到“5”(完全能接受)进行 5 点利克特量表评分。

(3)**对能够减少作业作弊行为策略的有效性评价(简称“策略有效性”)**

根据访谈结果，列举了旨在减少作业作弊行为的 9 种(Cronbach's α 系数为 0.781，具有较高的内部一致性)比较常用的策略(如“老师批作业更加仔细，从而找出那些作弊的同学”等)，请作答者就每一策略的有效性，从“1”(一点儿也没有用)到“5”(非常有用)进行 5 点利克特量表评分。

(4)**观察到的同伴的作业作弊行为(简称“同伴作弊行为”)**

采用三个条目进行测量：a. 同伴作业作弊行为的普遍性(简称“*同伴作弊–普遍性*”)，请作答者就作业作弊行为在周围同学中的普遍性进行“1”(从未见有人这样做过)到“5”(几乎所有人都会做)的 5 点利克特量表评分。

b. 同伴作业作弊行为的总体频率(简称“*同伴作弊–总体频率*”)，请作答者就周围同学做出作业作弊行为的频率进行“1”(从未)到“5”(经常)的 5 点利克特量表评分。

c. 同伴所采取的不同形式的作业作弊行为的频率(简称“*同伴作弊–具体频率*”)，列举了访谈中提及率较高的三种作业作弊的具体形式(即做作业时抄答案、抄他人作业、让他人

帮忙写作业)，请作答者就周围同学做出每一种具体作弊行为的频率，进行“1”(从未)到“5”(经常)的 5 点利克特量表评分。

(5)人口学信息

具体包括：学校类型(公办小学、民工子弟小学及民办小学)、年龄、性别、年级(2 至 6 年级)、是否为独生子女以及有关学习成绩的自我评价(自身成绩高于、等于或低于班级平均水平)。

2.4 机器学习模型建构

采用 SPSS 24.0 软件分析数据。首先进行描述性统计分析，在此基础上运用机器学习进行统计建模分析。考虑到机器学习算法的多样性，以及预测变量和结果变量间关系的未知性，本研究采用集成学习法，首先通过四种不同算法对小学生作业作弊行为的预测模型进行拟合，然后基于 Stacking 法将四种不同算法的训练结果进行整合。这四种算法具体为：逻辑回归、多层感知机(MLP)、极端梯度提升(XGBoost)和随机森林。

2.4.1 四种机器学习算法及步骤

逻辑回归是一种广义线性回归，是用逻辑函数解决对分类变量(以二分变量居多)进行预测的问题。多层感知机(MLP)是一种前馈人工神经网络模型，它由输入层、输出层和一个以上的隐藏层组成，每一层均有若干神经元，通过逐步调整神经元间不同连接的权重来拟合模型。极端梯度提升(XGBoost)是集成学习方法中常用的策略，通过不断训练及优化决策树，并将每次训练所得的输出值相加，以获得最终预测值。随机森林算法与 XGBoost 具有一定的相似性，但两者最大的区别在于随机森林的输出值仅仅是简单多数投票的结果，对不同决策树的训练结果也没有做进一步的优化提升。

本研究在 MLP 中构建的网络由隐藏层和双曲正切激活函数组成。在训练开始前首先将协变量进行标准化处理，训练过程中则采用共轭梯度下降法(Conjugate Gradient Descent, λ

= 0.0000005, σ = 0.00005, 区间中心 0, 区间偏移±0.5)来调整神经元间连接的权重，以达到将连续训练过程中的预测误差(针对划分至训练集的样本)最小化的目的。该算法还将在每次训练结束时计算测试集中的预测误差，以确保每次训练后预测误差的减少不是由于对训练集数据的过度拟合所致。XGBoost 是将训练过程中的所有模型参数设置为默认，并采用基于树模型(gbtree)的方法进行决策树递增。其中，决策树的递增数量为 100 棵，η(提升学习速率, boosting learning rate)默认值为 0.3，以其作为特征缩减技术(shrinkage)的设定值，同样用于防止训练集数据的过度拟合。此外，用于训练集和测试集的协变量都需预先进行标准化处理。随机森林算法在训练中构建决策树时采用的是自展法(bootstrap)抽样，决策树的数量同样默认为 100 棵，且寻找最佳分割属性时，把需要考虑的特征数量设置为样本特征的平方根。上述两个参数可用于提高预测准确性并防止模型过度拟合。此外，用于训练集和测试集的协变量也预先得到了标准化处理。

上述四种机器学习算法均包含以下五个步骤。第一，将所有数据随机拆分成三个独立的数据集：训练集、测试集和留出集(训练集占总被试量的 64%，测验集占 16%，留出集占 20%)。第二，在训练集中对数据进行训练，训练时共有 32 个特征输入。其中 22 个特征来自 9 道问卷问题(为表 2 的 8 个问题及其选项，以及学生对学习成绩的自我评价)；剩余 10 个特征是通过对学校、年级、性别及是否为独生子女等二分变量转换为虚拟变量(哑变量，其中学校的参照水平为普通公办学校，年级的参照为二年级，性别的参照为女，是否为独生子女参照独生子女)而得到。最终输入 32 个特征，通过训练获得第一个模型。第三，将该模型在测试集中进行测试，以获得其性能指标。第四，将训练集和测试集结合起来，重新将这些数据随机划分为新的训练集和测试集，重复第二和第三步，再次训练得到第二个模型。通过重复 100 次第二至第四步的建模过程(“划分–训练–测验–重组–划分”)，最终可获得 100 个模拟模型，从而确保无论样本被划分到训练集还是测试集，所得结果均能保持

一定程度的稳定性。第五，使用最初分配到留出集的数据(此前未参与过任何一次训练或测试)对模型进行最终验证，即进行外部效度的检验。

2.4.2 机器学习结果的集成

在采用上述四种机器学习算法进行分析的基础上，本研究使用 stacking 法对四种算法进行整合(即集成)。具体说来，首先设置逻辑回归、MLP、XGBoost 和随机森林四种算法对原始数据进行训练，然后将四种算法的训练结果进行整合(stacking 法设置为五折交叉验证法)，最后再对整合后的训练结果进行测试集验证和留出集验证(验证算法设置为逻辑回归)，从而实现对四种算法的取长补短，获得最终且最佳模型。

2.4.3 机器学习模型的关键指标

无论是针对每一种机器学习算法，还是对不同算法结果的集成，采用留出集对最终模型进行验证均可得到模型的两个关键指标。其一为计算机模型的受试者工作特征曲线(Receiver Operating Characteristic, ROC)下方面积(Area under the Curve, AUC)。ROC 和 AUC 是常用于衡量机器学习性能的指标，表现了分类模型的特征。其中，ROC 是通过真阳率(True Positive Rate, 正确预测出的阳性的数量/所有阳性的数量)和假阳率(False Positive Rate, 将阴性误判为阳性的数量/所有阴性的数量)绘制而来的曲线(见图 1)，AUC 值为 ROC 曲线下方的面积，AUC 越大表明分类器的效果越好，它可以用于评估模型的敏感性(即真阳率，本研究中表现为能够准确预测“存在作业作弊行为”这一情况的比例)和特异性(1–假阳率，本研究中表现为能够准确预测“不存在作业作弊行为”这一情况的比例)。敏感性和特异性之间存在着一种函数关系，在不同敏感性条件下，特异性会朝相反的方向变化，而在 ROC 曲线的对角线位置将达到两者的最优组合。在实际应用过程中，因具体应用情境不同，在 ROC 曲线上以多少的敏感性和特异性作为标准，需要根据实际情况来决定。

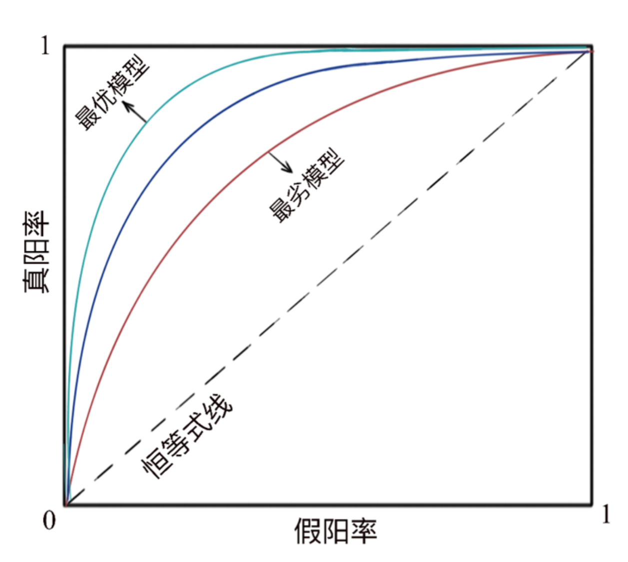


图 1 计算机模型的受试者工作特征曲线(ROC)描绘了模型敏感性(真阳率)随着特异性(假阳率)变化的规律。曲线下的面积(AUC)表示模型的整体性能，曲线距离恒等式线(虚线对角线)越远，代表模型的分类能力越好，反之则越差。

其二为各个预测变量的 Shapley 值。Shapley 值常用于处理复杂分配问题(Shapley, 1953)。如 A、B、C 三人共同完成一项工作，在为 A 分配奖金时，若要寻求相对公平，则需通过 Shapley 值法获得 A 在其中的边际贡献。该方法首先需要计算：仅有 A 参与时能够完成的工作量；B 和 C 与 A 合作时所能完成的工作量减去 B/C 单人能够完成的工作量；B 和 C 与 A 合作时完成的工作量减去仅 B 和 C 合作时完成的工作量。然后计算这些数值的均值，即为 A 最终的边际贡献。Shapley 值法全面考虑了在所有合作或单人关系中，A 所能带来的增益，对 A 的边际贡献性进行了公平且科学的计算。正因为此，越来越多的研究者开始借助 Shapley 值法，对预测变量在预测模型中的相对贡献性展开科学评估(Ghorbani & Zou, 2019; Lundberg & Lee, 2017)。

本研究的机器学习中，对各预测变量(影响因素)Shapley 值的计算完全遵循上述计算规则。其中，“A、B、C”即为不同的预测变量，而“工作量”为变量在模型中的贡献性大小，即预测准确性。预测变量的 Shapley 值代表其对总体模型预测准确性的实际边际贡献。通过计算每一个预测变量的 Shapley 值，Shapley 值的大小代表这些预测变量在模型中的相

对重要性大小。

3 结果与分析

3.1 描述性统计

3.1.1 作弊行为

图 2 显示了 2 至 6 年级小学生自我报告参与过作弊的人数百分比。由图可见，作弊行为的发生率在 3 年级时有了明显的增长，而在 4 年级开始相对趋于稳定。以年级为预测变量(其中设 2 年级为参考组)，作弊行为为结果变量，进行二元逻辑回归分析。结果表明，不同年级小学生的作弊率存在显著差异($p < 0.001$)，其中 2 年级的作弊率显著低于其他各年级(2 年级 vs 3 年级: $p = 0.002$, $B = 0.51$, $OR = 1.67$, $95\%\ CI = 1.21–2.29$; 2 年级 vs 4/5/6 年级: $ps < 0.001$, $B = 0.70, 0.57, 0.87$; $OR = 2.01, 1.77, 2.39$, $95\%\ CI = 1.46–2.75, 1.30–2.40, 1.77–3.22$)，但 3、4、5、6 四个年级间的作弊率两两差异不显著($ps > 0.05$)。

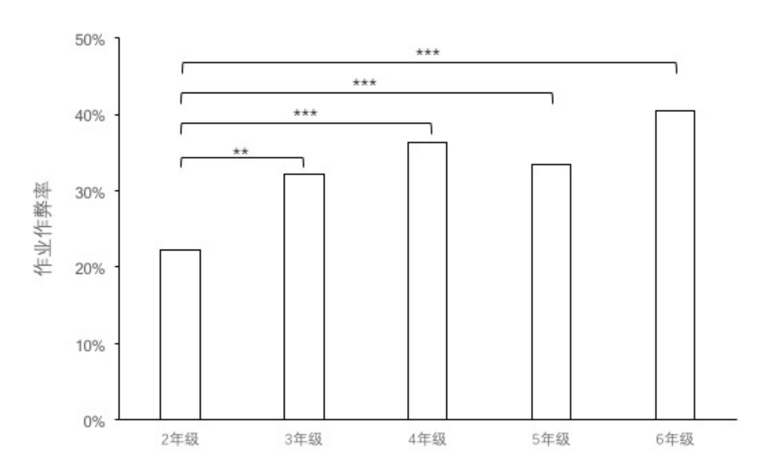


图 2 2–6 年级自我报告存在作弊行为的人数百分比(**$p < 0.005$, ***$p < 0.001$)

3.1.2 作弊行为的影响因素

表 2 显示了全体受试者在每个预测变量上的得分均值和标准差。由表可见，小学生普普遍认为“受老师惩罚”($M = 3.65$)是作弊的最严重后果，且作弊是不能被接受的($M = 2.02$，与 3 代表“中立”相比，$p < 0.001$)。与此同时，他们也倾向于认为周围同伴对作弊的接受程度较低($M = 2.08$，与 3 相比，$p < 0.001$)。此外，小学生普遍认为“加强课堂练

习，在课堂上把知识弄懂”($M = 3.28$)是减少作弊的最有效策略，“抄他人的作业”($M = 2.01$)是同伴作弊最常采用的形式。

表 2 作弊行为预测变量的描述统计结果

Item	Mean	SD
Q2.后果严重性, 1–5		
1. 受老师批评	3.05	1.22
2. 受老师惩罚	3.65	1.31
3. 受父母批评	3.24	1.25
4. 受父母惩罚	3.36	1.30
5. 受周围同学的批评或嘲笑	3.33	1.48
Q3.自我可接受性	1.81	1.12
Q4.同伴可接受性	2.13	1.17
Q5.策略有效性, 1–9		
1. 增加作弊后的惩罚力度	2.89	1.45
2. 老师批改作业更仔细	2.88	1.36
3. 加强课堂练习，在课堂上把知识弄懂	3.28	1.50
4. 老师加强教育，多强调作业作弊属于不良行为	2.47	1.39
5. 降低作业难度	2.12	1.39
6. 老师批评或惩罚	2.98	1.34
7. 给予独立完成作业的同学以表扬和奖励	2.88	1.42
8. 家长批评或惩罚	2.93	1.43
9. 做作业时父母监督	2.84	1.45
Q6.同伴作弊–普遍性	2.02	0.88
Q7.同伴作弊–总体频率	2.08	0.96
Q8.同伴作弊–具体频率, 1–3		
1. 抄他人的作业	2.01	1.05
2. 做作业时抄答案	1.91	1.04
3. 让他人帮写作业	1.24	0.61

3.2 机器学习分析

为了更加准确地计算不同变量对小学生作业作弊行为的预测效果，研究仅将预测变量相关问题作答率达到 100%的被试数据纳入机器学习，故最终进入机器学习分析的被试量

为 1637。如方法部分所述，本研究在使用四种机器学习算法进行分析的基础上，重点对四种算法所获结果进行集成。

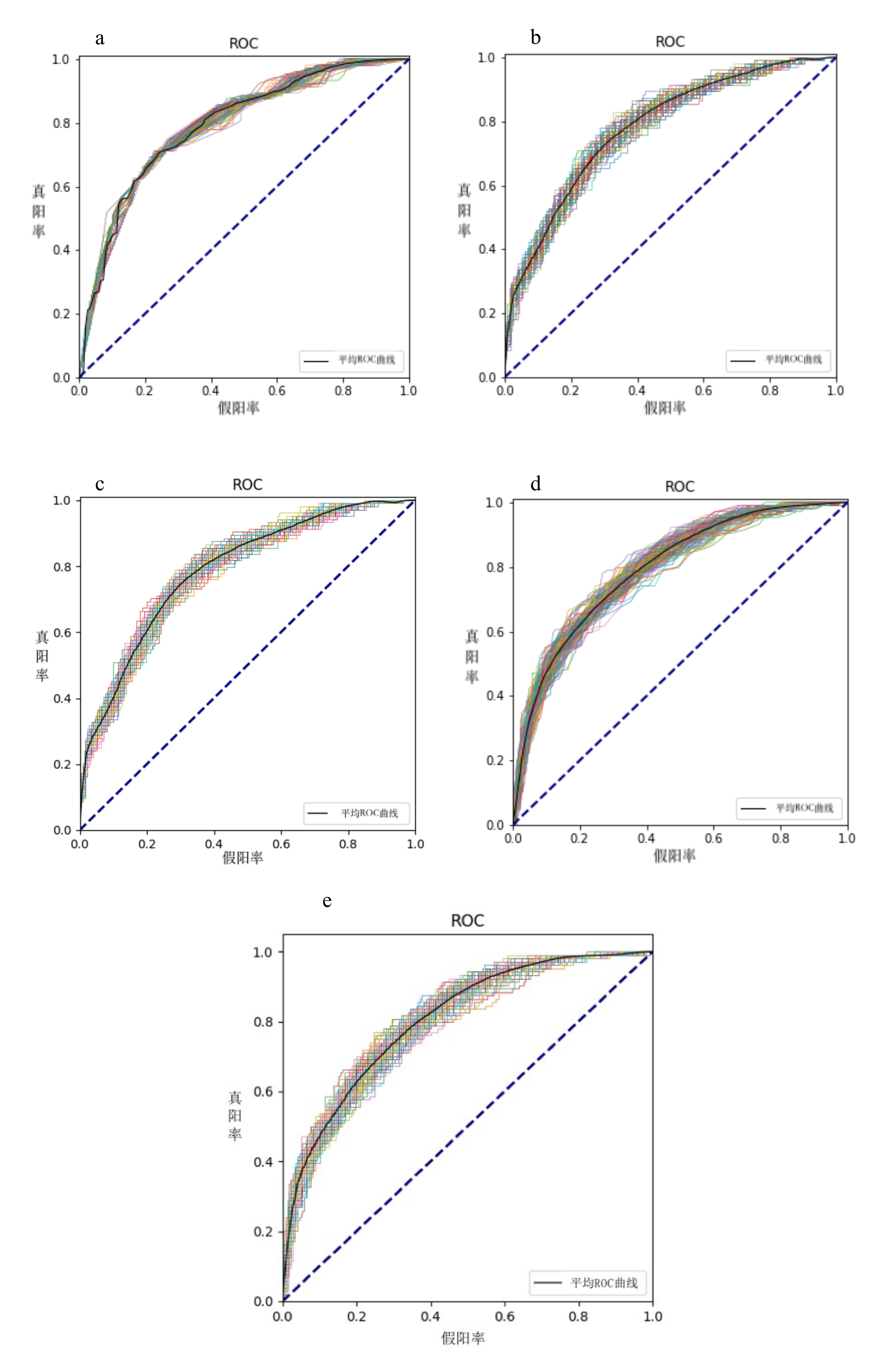
3.2.1 四种机器学习算法及集成学习法所获模型的 AUC 均值分析

表 3 显示了四种机器学习算法所各自模拟出的 100 个模型的 AUC 的均值，以及集成学习所模拟出的 100 个最终模型的 AUC 的均值。由表可见，这些 AUC 均值均显著高于随机水平(50%, $ps < 0.001$)。其中，集成学习法的最终预测模型的 AUC 均值达到了 80.46%，意味着该模型有 80.46%的概率能够准确预测小学生是否会做出作业作弊行为。

表 3 四种机器学习算法及集成学习法下留出集验证所产生的 100 个模型的 AUC(%)均值和标准差

模型	均值(%)	标准差	95% 置信区间	
			Lower	Upper
逻辑回归	77.87	1.50	77.72	78.01
XGBoost	77.82	1.69	77.63	78.02
MLP	78.25	1.70	78.01	78.48
随机森林	79.47	0.95	79.28	79.66
集成学习	**80.46**	**0.80**	**80.30**	**80.62**

图 3 显示了四种不同算法以及集成学习分别所得的 100 个模型的具体 AUC 值(见图 3)。如图所示，集成学习算法下的模型总体具有较高的敏感性和特异性(1–假阳性率)。由于集成学习法综合了四种算法的优点，故后续的分析将重点关注集成学习法所获结果。将表 3 中集成学习法的 AUC 均值转换为 Cohen’s *d* 值，所得的 Cohen’s *d* 均值为 1.214(95%置信区间: 1.205–1.222)，表明该方法所获的 100 个最终模型的平均效应量很大(Cohen, 1988, Cohen’s $d > 1.2$ 时表示效应量非常大)。



注：a、b、c、d、e 分别为逻辑回归、XGBoost、MLP、随机森林及集成学习算法的 AUC 值。

图 3 四种不同算法及集成学习分别获得的 100 个模型的 AUC 值

3.2.2 集成学习法所获最终模型中各变量的 Shapley 值分析

在集成学习中，采用留出集对最终模型进行检验，得到所有预测变量的 Shapley 均值，这些值均显著高于 0 ($ps < 0.05$)，即所有预测变量皆对该模型具有显著的边际贡献。将这些预测变量按照 Shapley 值的大小进行排列，可直观显示这些预测变量在模型中的相对重要性排序。图 4 列举了对模型预测准确性的边际贡献大于或等于 1%的主要预测变量。这些变量间的相对重要性依然存在较大的差异。根据重要性由高到低可大致将这些预测变量分为以下四个组。

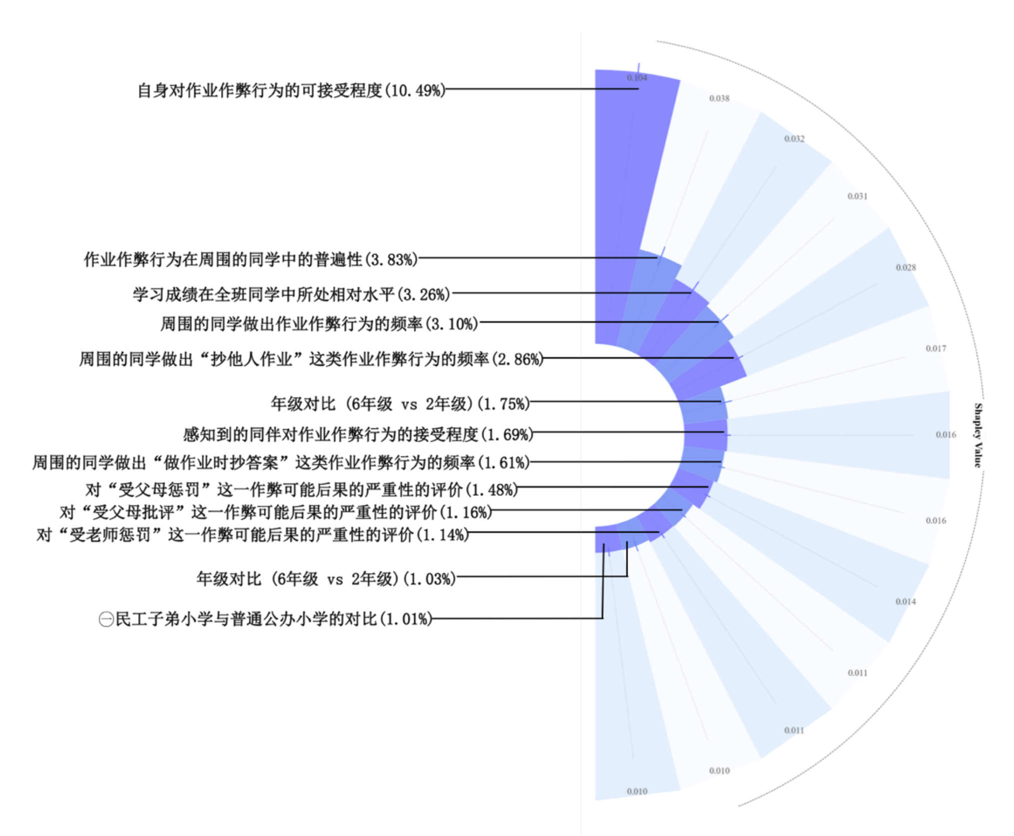


图 4 主要预测变量在预测作业作弊行为时的 Shapley 值及其 95%置信区间。其中“㊀”表示该预测变量对结果变量存在负向预测作用；其他预测变量对结果变量均存在正向预测作用。

第一组为小学生自身对作弊的接受性(接受性越高，就越有可能作弊)。该预测变量的 Shapley 值高达 10.49%，代表其对总体模型预测准确性的实际边际贡献为 10.49%。配对样本 t 检验结果显示，该变量的 Shapley 值显著高于 3.83%(排序第二的预测变量对应 Shapley 值, $t = 23.88$, $df = 327$, $p < 0.001$)，表明其在模型中的预测力最佳且远高于其他变量。

第二组变量的 Shapley 值均在 2%–4%内。首先是同伴作弊的普遍性，该预测变量的 Shapley 值为 3.83%，对模型预测准确性的贡献性显著高于剩余所有变量(vs 3.26%, t = 1.98, df = 327, p = 0.048)。其次是小学生自身成绩在全班同学中所处的相对水平(反向计分)、同伴作弊行为的普遍性、同伴作弊的频率，以及同伴做出“抄他人作业”这一特定作弊行为的频率。这三个变量的 Shapley 值间不存在显著的差异(具体数值参见图 4, t = 0.57, 1.78, 1.23, dfs = 327, p = 0.57, 0.22, 0.08)，且后两者间有显著的相关关系(r = 0.21, p < 0.001)，即三者对小学生作业作弊行为的预测力相当。具体说来，上述变量均对作弊行为存在显著的正向预测作用，即自我评价的成绩越差也越有可能作弊；同伴作弊行为发生得越普遍或越频繁，个体自身越有可能作弊。

第三组变量的 Shapley 值均在 1%–2%内，在模型中的边际贡献较小，该组变量中最大的 Shapley 值(1.75%)显著低于 2.9%(t = – 6.99, df = 327, p < 0.001)。包括：小学生所处年级(6 和 4 年级学生的作弊率高于 2 年级学生)，感知到的同伴对作弊行为的接受程度，同伴做出“做作业时抄答案”这一作弊行为的频率(感知到的可接受性/频率越高，作弊的可能性越大)，小学生对“受父母惩罚”、“受父母批评”和“受老师惩罚”这三种作弊可能带来的后果的严重性评价(越是认为这些后果严重，越不可能作弊; r = 0.44, 0.34, 0.36, ps < 0.001)，以及学校类型(民工子弟学校的作弊率低于普通公办学校)。

除了上述预测变量，剩余的变量虽对模型预测准确性具有显著的贡献(ps < 0.05)，但重要性非常弱(Shapley 值< 0.01)。这些预测力非常弱的变量主要包括：小学生对所有 9 项减少作弊的策略的有效性评价，对部分作弊可能带来后果的严重性评价(“受老师批评”和“受周围同学的批评或嘲笑”)，以及人口学信息(见表 4)。

表 4 次要预测变量在预测作业作弊行为时的 Shapley 值的均值(%)和标准差及其 95%置信区间

预测变量	均值 (%)	标准差	95% 置信区间	
			Lower	Upper
民办小学与普通公办小学的对比	0.91	1.38	0.76	1.06
对“受周围同学的批评或嘲笑”这一作弊可能后果的严重性的评价	0.87	0.71	0.79	0.95
年级对比（3 年级 vs 2 年级）	0.80	1.84	0.61	1.00
对“降低作业难度”这一减少作弊的策略的有效性评价	0.70	0.80	0.62	0.79
对“给予独立完成作业的同学以表扬和奖励”这一减少作弊的策略的有效性评价	0.68	0.60	0.61	0.74
对“做作业时父母监督”这一减少作弊的策略的有效性评价	0.63	0.59	0.56	0.69
对“老师加强教育，多强调作业作弊属于不良行为”这一减少作弊的策略的有效性评价	0.59	0.60	0.53	0.66
对“老师批改作业更仔细”这一减少作弊的策略的有效性评价	0.58	0.52	0.52	0.64
对“受老师批评”这一作弊可能后果的严重性的评价	0.54	0.58	0.48	0.61
年级对比（5 年级 vs 2 年级）	0.49	0.91	0.39	0.59
对“增加作弊后的惩罚力度”这一减少作弊的策略的有效性评价	0.47	0.52	0.42	0.53
对“老师批评或惩罚”这一减少作弊的策略的有效性评价	0.47	0.46	0.42	0.52
对“加强课堂练习，在课堂上把知识弄懂”这一减少作弊的策略的有效性评价	0.46	0.49	0.41	0.52
对“家长批评或惩罚”这一减少作弊的策略的有效性评价	0.45	0.38	0.41	0.49
性别对比（女生 vs 男生）	0.40	0.36	0.36	0.44
有弟弟/妹妹者与独生子女的对比	0.38	0.71	0.30	0.45
有哥哥/姐姐者与独生子女的对比	0.22	0.55	0.16	0.28
周围同学让他人帮写作业这一行为的频繁性	0.13	0.41	0.09	0.18
既有哥哥/姐姐又有弟弟/妹妹者与独生子女的对比	0.06	0.34	0.02	0.10

4 讨论

本研究采用问卷调查及机器学习这一人工智能的核心方法，首次系统考察了小学生的作业作弊行为，及其与个体认知和情境变量(对作弊后果的严重性的感知、自身对作弊的接

受性和感知到的同伴对作弊的接受性、对作弊干预策略的有效性评价等)、同伴作弊的普遍性和频率，以及性别、年级及学校类型等人口学变量之间的关系，得到以下结果。

第一，总体而言，约 33%的小学生自我报告曾经有过作业作弊行为，且这一作弊发生率随年级的增长呈现出一定的变化趋势。具体表现为：2 年级为小学生作业作弊的萌发阶段；3 年级为激增阶段，随后进入平缓期；到 6 年级时，小学生作业作弊率已达 40.5%。这表明，作业作弊在 2 年级小学生中就已经出现，到 6 年级已相对普遍。导致 3 年级作业作弊率迅速增长的原因可能在于 3 年级的作业量及作业频率的增加。而 4 年级及以后表现出的作弊率趋于平缓，这表明了道德发展水平在小学阶段并非影响其作弊行为的关键因素(该假设在小学生说谎行为的研究中得到了证实，见 Lee, 2013)。由此可见，作弊作为一种习惯性行为，一旦形成，若不及时采取有效的干预手段，将可能持续存在甚至呈增长趋势，这从侧面也反映出早期诚信教育和对作弊进行早期干预的重要性。

第二，本研究所考察的各个预测变量均对小学生作业作弊行为具有显著的预测效果。机器学习结果表明，通过集成学习法对不同算法所获模型进行整合，最终模型在预测作业作弊行为时具有较高的敏感性和特异性(AUC 均值高达 80.46%)。这一结果意味着：该模型有 80.46%的可能性能够基于这些预测变量对小学生是否作弊作出正确预测。转化 AUC 值所得的 Cohen's *d* 也显示，模型在通过这些预测变量对小学生作业作弊行为进行预测时，所得结果的效应量很大(Cohen's $d > 1.2$)。

第三，机器学习模型各变量的 Shapley 值分析结果进一步表明，所有预测变量均对最终模型的较高预测准确率和良好拟合作出了显著贡献。然而，这些预测变量的重要性存在比较明显的差异。具体说来：

首先，小学生自身对作业作弊行为的接受程度是其是否作弊的最关键的预测变量。他们越是认为作业作弊是可接受的，就越有可能报告自己曾经有过这样的行为。这一结果与

以往在大中学生群体中的发现基本一致(Abaraogu et la., 2016; Ives & Giukin, 2020)。Murdock 和 Anderman 通过对以往研究进行综述后提出，个体对作弊行为可接受性的判断差异可能基于两种原因：第一，个体认为作弊的可接受性高可能是由于其并不认为作弊是不道德的行为；第二，个体虽然知道作弊是不道德的行为，但出于为自己辩护，故对其自身作弊行为进行了合理化，以此减少作弊行为与自身道德认知的冲突(Murdock & Anderman, 2006)。由此可见，与大中学生一样，小学生对作弊行为可接受性的判断可能也受到了这两个原因的影响。

其次，同伴做出作业作弊行为的普遍性和频率也是个体是否作弊的重要预测变量，即同伴作业作弊行为是影响小学生自身作业作弊行为的重要因素。该结果与以往针对大学生和中学生群体的研究所获结果较为一致，即同伴作弊行为能够导致个体对其自身作弊行为进行合理化(Hrabak et al., 2004; Ghanem & Mozahem, 2019; McCabe & Treviño, 1993)。近期一项针对来自世界多国研究的元分析也显示，同伴作弊是个体学业作弊行为的最重要因素之一，且这种“跟风效应”受到文化因素的调节——其在集体主义文化背景下较个体主义文化背景表现更为明显(Zhao & Mao et al., 2022)。

同时，对学业成绩的自我评价这一预测变量也具有与同伴作弊行为相似的重要性。具体表现为，自我评价成绩较差的小学生更有可能报告自己有过作业作弊行为。这一结果也与针对大中学生的研究所获结果基本相同(Newstead et al., 1996; Özcan et al., 2019)，其反映出作弊行为背后的一种潜在动机：成绩较差的学生希望通过作弊来避免失败(参见 Oran et al., 2016)。此外，还有研究发现，成绩好的学生更容易受到“学习、掌握新知识”这一学习动机的驱动，成绩较差的学生却更容易受到“获得较高的成绩”的驱动，为了考得更好而做出作弊行为(Putarek & Pavlin–Bernardic, 2020)。还有一种可能性是，自我评价学业成绩较差的学生的自我效能感可能相对较弱，而自我效能感弱的学生往往更容易做出学业作弊

行为(see Murdock & Anderman, 2006)。

此外，对小学生作业作弊行为具有一定预测作用的因素还包括年级，学校类型，同伴对作弊的接受性和作弊后果的严重性，但这些变量的预测效应相对较弱。例如，本研究发现，来自普通公办小学的学生相较于民工子弟小学的学生，表现出了更高的作业作弊率。导致这种学校间差异的原因可能与校园风气、教学政策以及师生关系等有关(Ramberg & Modin, 2019; McCabe et al., 2012)，这些因素往往决定了一所学校所特有的学习氛围。

再如，本研究发现，感知到的同伴对作弊行为的接受程度对小学生作业作弊行为具有较弱的预测作用。并且，该变量的预测力明显弱于小学生自身对作弊的可接受性，及观察到同伴的作弊行为的普遍性或频繁性。由此可见，同伴行为对作弊的影响主要表现在其被观察到的外部行为上，即小学生通过观察同伴是否作弊，达到对自身行为进行合理化的目的(Jurdi et al., 2011; McCabe & Abdallah, 2008)。然而，他人的道德意识水平并非是影响小学生是否作弊的主要因素。研究发现“受周围同学的批评或嘲笑”这一作弊后果严重性对作弊的预测作用也非常小，该结果在一定程度上佐证了上述结论。

对作弊后果的严重性感知也是较为次要的预测因素，即作弊后果严重性对小学生作业作弊行为的预测贡献性较小。这一结果与本研究假设相悖。究其原因，可能是因为我国并不存在实质性的作业作弊“惩罚”制度。小学生对作业作弊后果的认知非常有限，有限的认知主要来源于对考试作弊可能导致后果的认识(如被父母、老师惩罚等)。但值得反思的是，尽管各级各类学校大多制定了严格的考试作弊“惩罚”制度，却始终无法彻底杜绝考试作弊行为。这也表明，过度依赖于外部强制力量而忽略了对学生自身能动性的培养(即对学生道德价值观的培养)的作弊惩罚方法是无法真正起到遏制作弊的效果。

最后，小学生对减少作弊策略有效性的评价及其他部分人口统计学因素是重要性最低的预测因素(但其对模型预测的重要性仍然显著高于 0)。其中，就减少作弊策略的有效性而

言，之所以其预测作用弱，很可能与小学生普遍倾向于认为这些常见的抑制作业作弊行为的策略效果不佳、对其有效性打分普遍偏低有关。

本研究具有一定的理论创新性：作为国内首个聚焦于小学生作业作弊行为及其影响因素的研究，初步建立了关于小学生作业作弊影响因素的综合模型，并将这些影响因素根据量化后的重要性指标进行了排序。特别需要指出的是，本研究发现小学生作业作弊的影响因素与大中学生学业作弊行为的影响因素构成并不完全相同，存在其特殊性。如在Murdock 和 Anderman (2006)的模型中，后果严重性是能够预测大中学生考试作弊的至关重要的作弊动机，但该指标对小学生作业作弊的影响却远不如预期之大。可见小学生群体具有其发展的特殊性。此外，本研究还发现，除了后果严重性，还有一些被以往研究证明能够影响大学生和中学生学业作弊行为的因素，在预测小学生作业作弊时贡献也非常微弱，如是否是独生子女(Tsui & Ngo, 2016, 兄弟姐妹数量越多的大学生越有可能作弊)、性别(Gallowaya, 2012, 男高中生作弊率高于女生)等；但部分对小学生作业作弊具有较强预测作用的变量在大中学生学业作弊行为中却不存在显著效应，如年级(Ives et al., 2017)等。

但是，研究结果也表明， Murdock 和 Anderman (2006)提出的一些影响大中学生作弊行为的关键要素也能够预测小学生的作业作弊，即也存在一定的相似性。如个体自身对作业作弊行为的可接受性，同伴作业作弊的普遍性和发生频率，及其自身的成绩水平是预测小学生作业作弊的关键变量。可见，多个影响着大学生和中学生学业作弊行为的因素，也同样会对小学生的作业作弊行为存在显著影响(Abaraogu et al., 2016; Ghanem & Mozahem, 2019)。

此外，本研究还发现，个体对情境的认知(如“个体自身对作业作弊行为的可接受性”等)在影响作业作弊行为中也发挥着重要的作用。这一结果为一个长久以来的有关作弊是情境驱动还是个体特质驱动这一争议提供了新的思路和理论依据。过去关于作弊的研究

似乎更加认可个体的作弊行为是受情境而非个人特质驱动的(Hartshorne & May, 1928)。而本研究的结果则表明，情境和个人特质的交互作用可能也在其中发挥着重要的作用。

本研究在方法上也具有一定的创新性。近年来，随着人工智能在医学、化学及生物学等领域的广泛应用，机器学习技术已在成人(Bartlett et al., 2014; Just et al., 2017; Livieris et al., 2018)乃至儿童(Bruer et al., 2019; Zanette et al., 2016)研究中得到越来越广泛的应用。本研究首次将机器学习运用于儿童道德研究领域，对小学生作业作弊行为的影响因素进行深入分析。结果表明，机器学习法在分析儿童行为发展数据上是可行的，在数智时代背景下，为后续更多的儿童心理学领域同行提供了一种新的研究和数据分析思路。

本研究还具有较强的实践创新性。本研究采用机器学习方法创建了能够预测小学生有多大的概率会做出作业作弊行为的模型(预测准确率达 80%以上)，未来可基于该模型创建相关 app 或网页版测试，在小学中投入应用。具体来说，首先由小学生在 app 或网页上填写本研究的问卷并提交，后台基于已有模型对提交的数据进行计算并输出结果(即每位小学生作业作弊的可能性或概率, 0%至 100%)。根据输出的结果，教师或家长能够初步了解小学生当下有多大的可能性会在作业中作弊，以便于教师或家长进行必要的关注。但值得注意的是，我们并不能依据一次测试的结果就给学生“贴标签”，关于如何使用预测模型的结果，还需结合研究伦理和学校教育方式加以综合考虑。此外，由于目前问卷涉及的部分预测变量对预测作业作弊的贡献性非常小，如对预防作弊的方法的有效性评价等。后续研究可继续删除问卷中的相关条目并重新建模，开发出更简短的问卷。此举将进一步推进研究成果在学校和家庭教育中的推广应用。

此外，本研究还为诚信教育的方式方法提供了具体、可操作的实践方案。首先，个体对作弊的可接受性(或者说“道德意识”)，而非作业作弊后果的严重性，能够较好地预测小学生的作业作弊行为。这就给予家长、老师及广大教育工作者以启示：不应一味强调作

弊的后果，相反，更有效的做法是从积极的角度帮助学生建立正确的学业诚信认知，引导其认识到作弊是“零容忍”的不道德行为。其次，由于同伴作弊对个体作业作弊影响很大，因此教师和家长应注意减少同伴行为所产生的负面影响，并尝试通过各种方法树立同伴的正面影响。例如，给予那些即使作业做得不好，但仍然选择独立完成的学生以充分肯定，让学生认识到，独立完成作业比正确率更重要(Misselbrook, 2014; Siev & Kliger, 2019)。再次，学校还应积极响应教育部倡导的“双减”政策，重“质”不重“量”，并着力建立起一种良性的学习氛围，杜绝“唯成绩论”：让学生认识到，做作业仅仅是对课上知识的查漏补缺和及时巩固，并不是为了让同班同学“一较高下”(Misselbrook, 2014; Siev & Kliger, 2019)。最后，作弊是一种习惯性行为(Davy et al., 2007)，鉴于小学生的学业作弊率在所有年龄段中相对较低(Cochran, 2015)，尚处于萌发阶段，我们更应该“从娃娃抓起”，适时在小学乃至幼儿园阶段推广诚信教育，将作弊行为“扼杀在摇篮里”。

本研究存在一定的局限性。首先，本研究所创建的小学生作业作弊行为的预测模型达到了 80.46%，但仍存在近 20%的预测错误的可能性，因此仍需在未来的研究中进一步优化模型。一方面，可以通过理论探索，寻找更多新的可能影响作业作弊的预测变量来完善模型。本研究主要参考了影响大学生和中学生学业作弊行为的因素，且侧重于考察与学校有关的情境因素对小学生作业作弊行为的影响，因此不能排除还有其他预测因素的存在。且研究结果表明，相较于大中学生，影响小学生作业作弊的因素也具有一定的特殊性，未来的研究可依据上述领域进行拓展，更充分地挖掘影响小学生作弊的因素，进而创建一个科学而完备的、专门适用于小学生作业作弊的理论模型。另一方面，机器学习是服务于大数据的方法，未来可将本研究问卷转变为 app 或网页，通过进一步扩大样本量可以通过进一步扩大样本量，再增加一个外部验证的数据集，对参数进行调优，进一步完善模型，提高模型的预测准确性。其次，本研究对于小学生作业作弊行为的测量采用的是自我报告法，

虽然问卷是匿名性质的，但受社会称许性影响仍存在个体自我报告的作弊率低于真实作弊率这一可能性。未来研究可尝试通过自我报告法和行为实验法相结合的方式来解决该问题。事实上，最近已有越来越多的研究者开始采用实验法来研究作弊问题，但这些实验范式比较适合年龄相对较大的学生群体或成人(Cizek, 1999; Zhao et al., 2021; Zhao & Peng et al., 2022)，实验法在儿童群体中的适用性还有待于更多研究加以检验。最后，本研究考察了不同作业作弊方式的普遍性，但这些方式以“抄作业”、“抄答案”等传统方式为主，并未涉及数智时代下一些新型的作弊方式(如，采用“爱作业”等批改作业的 app 进行作弊，或其他“线上”方式作弊)。未来研究可对上述新型作弊方式加以探讨。

5 结论

本研究首次聚焦小学生作业作弊行为，采用机器学习集成算法，系统考察了影响小学生作业作弊行为的关键因素及其相对重要性。结果表明，33%的小学生自我报告曾经做过作业作弊行为，且作弊行为的发生率随年级的增长呈上升趋势；机器学习集成算法得到的预测模型 AUC 值达 80.46%，能够较准确地预测小学生的作业作弊行为。总的来说，小学生的作业作弊行为很大程度上取决于他们对于作业作弊的接受性，同伴的作业作弊行为，及其自身的成绩水平。上述结果为儿童诚信行为发展的理论构建以及学业作弊的早期干预提供了科学依据。此外，本研究也表明，机器学习作为人工智能的核心方法，是一种可用于分析发展性数据分析的有效方法。

参考文献

Abaraogu, U. O., Henning, M. A., Okpare, M. C., & Rajput, V. (2016). Disclosing academic dishonesty: perspectives from Nigerian and New Zealand health professional students. *Ethics & Behavior*, *26*(5), 431–447. https://doi.org/10.1080/10508422.2015.1055494

Anderman, E. M., & Midgley, C. (2004). Changes in self–reported academic cheating across the transition from middle school to high school. *Contemporary Educational Psychology*, *29*(4), 499–517. https://doi.org/10.1016/j.cedpsych.2004.02.002

Bartlett, M. S., Littlewort, G., Frank, M., & Lee, K. (2014). Automatic decoding of facial movements reveals deceptive pain expressions. *Current Biology*, *24*(7), 738–743. https://doi.org/10.1016/j.cub.2014.02.009

Błachnio, A. (2019). Don't cheat, be happy. Self–control, self–beliefs, and satisfaction with life in academic honesty: A cross–sectional study in Poland. *Scandinavian Journal of Psychology*, *60*, 261–266. https://doi.org/10.1111/sjop.12534

Bong, M. (2008). Effects of parent–child relationships and classroom goal structures and motivation, help–seeking avoidance, and cheating. *Journal of Experimental Education*, *76*(2), 191–217. https://doi.org/10.3200/JEXE.76.2.191–217

Bretag, T., Mahmud, S., Wallace, M., Walker, R., McGowan, U., East, J., ... & James, C. (2014). 'Teach us how to do it properly!' An Australian academic integrity student survey. *Studies in higher education*, *39*(7), 1150–1169. https://doi.org/10.1080/03075079.2013.777406

Bruer, K. C., Zanette, S., Ding, X. P., Lyon, T. D., & Lee, K. (2019). Identifying liars through automatic decoding of children's facial expressions. *Child Development*, *91*, e995–e1011. https://doi.org/10.1111/cdev.13336

Bucciol, A., Cicognani, S., & Montinari, N. (2017). *Cheating in academia: The relevance of social factors*. Working Papers 15/2017, University of Verona, Department of Economics.

Campbell, D. T. (1986). Relabeling internal and external validity for applied social scientists. *New Directions for Program Evaluation*, *31*, 67–77. https://doi.org/10.1002/ev.1434

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd ed.)*, Hillsdale, NJ: Erlbaum.

Cochran, J. K. (2015). The effects of life domains, constraints, and motivations on academic dishonesty: a partial test and extension of Agnew's general theory. *International Journal of Offender Therapy and Comparative Criminology*, *61*(11), 1288. https://doi.org/10.1177/0306624X15618689

Cizek, G.J. (1999), *Handbook of educational policy*, Academic Press, San Diego, CA.

Davy, J. A., Kincaid, J. F., Smith, K. J., & Trawick, M. A. (2007). An examination of the role of attitudinal characteristics and motivation on the cheating behavior of business students. *Ethics & Behavior*, *17*(3), 281–302. https://doi.org/10.1080/10508420701519304

Freire, C. (2014). Academic misconduct among Portuguese economics and business undergraduate students–a comparative analysis with other major students. *Journal of Academic Ethics*, *12(*1), 43–63. https://doi.org/10.1007/s10805–013–9199–2

Galloway, M. K. (2012). Cheating in advantaged high schools: prevalence, justifications, and possibilities for change. *Ethics & Behavior*, *22*(5), 378–399. https://doi.org/10.1080/10508422.2012.679143

Ghanem, C. M., & Mozahem, N. A. (2019). A study of cheating beliefs, engagement, and perception–the case of business and engineering students. *Journal of Academic Ethics*, *17*, 291–312. https://doi.org/10.1007/s10805–019–9325–x

Cheung, H. Y., Wu, J., & Huang, Y. (2016). Why do Chinese students cheat? Initial findings based on the self–reports of high school students in China. *The Australian Educational Researcher*, *43*(2), 245–271. https://doi.org/10.1007/s13384–016–0201–z

Ghorbani, A., & Zou, J. (2019). *Data shapley: Equitable valuation of data for machine learning*. Working Paper

Golino, H. F., Gomes, C. M. A., & Andrade, D. (2014). Predicting academic achievement of high–school students using machine learning. *Psychology*, *5*(5), 2046–2057. https://doi.org/10.4236/psych.2014.518207

Gullifer, J. M., & Tyson, G. A. (2014). Who has read the policy on plagiarism? Unpacking students' understanding of plagiarism. *Studies in Higher Education*, *39*(7), 1202–1218. http://dx.doi.org/10.1080/03075079.2013.777412

Hartshorne, H. & May, M.A. (1928). *Studies in deceit.* New York: McMillan.

Hrabak, M., Vujaklija, A., Vodopivec, I., Hren, D., Marušić, M., & Marušić, A. (2004). Academic misconduct among medical students in a post–communist country. *Medical Education*, *38*, 276–285. https://doi.org/10.1111/j.1365–2923.2004.01766.x

Ives, B., Alama, M., Mosora, L. C., Mosora, M., Grosu–Radulescu, L., & Clinciu, A. I., et al. (2017). Patterns and predictors of academic dishonesty in Romanian university students. *Higher Education*, *74*, 815–831. https://doi.org/10.1007/s10734–016–0079–8

Ives, B., & Giukin, L. (2020). Patterns and predictors of academic dishonesty in Moldovan university students. *Journal of Academic Ethics*, *18*(1), 71–88. https://doi.org/10.1007/s10805–019–09347–z

Jurdi, R., Hage, H. S., & Chow, H. P. H. (2011). Academic dishonesty in the Canadian classroom: behaviours of a sample of university students. *Canadian Journal of Higher Education*, *41*(3), 1–35. https://doi.org/10.47678/cjhe.v41i3.2488

Just, M. A., Pan, L., Cherkassky, V. L., McMakin, D. L., Cha, C., Nock, M. K., & Brent, D. (2017). Machine learning of neural representations of suicide and emotion concepts identifies suicidal youth. *Nature Human Behaviour*, *1*(12), 911. https://doi.org/10.1038/s41562–017–0234–y

Kam, C. C. S., Hue, M., & Cheung H. Y. (2017). Plagiarism of Chinese secondary school students in Hong Kong. *Ethics & Behavior*, *28*(4), 316–335. https://doi.org/10.1080/10508422.2017.1333909

Lee, K. (2013). Little liars: development of verbal deception in children. *Child Development perspectives*, *7*(2), 91–96. https://doi.org/10.1111/cdep.12023

Lee, S. D., Kuncel, N. R., & Gau, J. (2020). Personality, attitude, and demographic correlates of academic dishonesty: a meta–analysis. *Psychological Bulletin*, *146*(11). https://doi.org/10.1037/bul0000300

Lim, V. K. G., & See, S. K. B. (2001). Attitudes toward, and intentions to report, academic cheating among students in Singapore. *Ethics & Behavior*, *11*(3), 261–274. https://doi.org/10.1207/S15327019EB1103_5

Livieris, I. E., Kotsilieris, T., Dimopoulos, I., & Pintelas, P. (2018). Decision support software for forecasting patient's length of stay. *Algorithms*, *11*(12), 199. https://doi.org/10.3390/a11120199

Lundberg, S. M., & Lee, S. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, 4765–4774. https://doi.org/10.48550/arXiv.1705.07874

McCabe, D. L., & Abdallah, F. H. (2008). Academic dishonesty in the Middle East: individual and contextual factors. *Research in Higher Education*, *49*(5), 451–467.

McCabe, D. L., Butterfield, K. D., & Trevino, L. K. (2012). *Cheating in college: why students do it and what educators can do about it*. Baltimore, MD: John Hopkins University Press.

McCabe, D. L., & Treviño, L. K. (1993). Academic dishonesty: honor codes and other contextual influences. *The Journal of Higher Education*, *64*, 522–538.

McCabe, D. L., & Treviño, L. K. (1997). Individual and contextual influences on academic dishonesty: a multicampus investigation. *Research in Higher Education*, *38*(3), 379–396. https://doi.org/10.1023/A:1024954224675

Misselbrook, D. (2014). Is honesty the best policy? why trustworthiness is no easy answer. *British Journal of General Practice the Journal of the Royal College of General Practitioners*, *64*(628), 558–559. https://doi.org/10.3399/bjgp14X682129

Molnar, K. K., & Kletke, M. G. (2012). Does the type of cheating influence undergraduate students' perceptions of cheating? *Journal of Academic Ethics*, *10*(3), 201–212. https://doi.org/10.1007/s10805–012–9164–5

Murdock, T. B., & Anderman, E. M. (2006). Motivational perspectives on student cheating: toward an integrated model of academic dishonesty. *Educational Psychologist*, *41*, 129–145. http://dx.doi.org/10.1207/s15326985ep4103_1

Newstead, S. E., Franklyn–Stokes, A., & Armstead, P. (1996). Individual differences in student cheating. *Journal of Educational Psychology*, *88*(2), 229–241.

Oran, N. T., Can, H. O., Senol, S., & Hadimli, A. P. (2016). Academic dishonesty among health science school students. *Nursing Ethics*, *23*(8), 919–931. https://doi.org/10.1177/0969733015583929

Özcan, M., Yeniçeri, N., & Çekiç, E. G. (2019). The impact of gender and academic achievement on the violation of academic integrity for medical faculty students, a descriptive cross–sectional survey study. *BMC Medical Education*, *19*:427. https://doi.org/10.1186/s12909–019–1865–7

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., & Grisel, O., et al. (2011). Scikit–learn: machine learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830. https://doi.org/10.1524/auto.2011.0951

Putarek, V., & Pavlin–Bernardic, N. (2020). The role of self–efficacy for self–regulated learning, achievement goals, and engagement in academic cheating. *European Journal of Psychology of Education*, *35*(3), 647–671. https://doi.org/10.1007/s10212–019–00443–7

Ramberg, J., & Modin, B. (2019). School effectiveness and student cheating: do students' grades and moral standards matter for this relationship? *Social Psychology of Education*, *22*(3), 517–538. https://doi.org/10.1007/s11218–019–09486–6

Shapley, L. S. (1953). A value for n–person games. *Contributions to the Theory of Games*, *2*, 307–317.

Shu, S., Sang, Q., Guo, Y., Huang, X. (2018). Why cheating Is contagious: the mediating effect of social loss and attitude toward cheating. Psychological Development and Education, 34(6), 664–671. https://doi.org/10.16187/j.cnki.issn1001–4918.2018.06.04

Siev, S., & Kliger, D. (2019). Cheating in academic exams: a field study. *Dishonesty in Behavioral Economics*, 111–140. https://doi.org/10.1016/B978–0–12–815857–9.00008–X

Smith, M., & Alvarez, F. (2021). Identifying mortality factors from machine learning using Shapley values – a case of COVID19. *Expert Systems with Applications*, *176*, 114832. https://doi.org/10.1016/j.eswa.2021.114832

Tsui, A. P. Y., & Ngo, H. Y. (2016). Social predictors of business student cheating behaviour in Chinese societies. *Journal of Academic Ethics*, *14*(4), 1–16. https://doi.org/10.1007/s10805–016–9266–6

Williams, M. W. M, & Williams, M. N. (2012). Academic dishonesty, self–control, and general criminality: a prospective and retrospective study of academic dishonesty in a New Zealand university. *Ethics & Behavior*, *22*(2), 89–112.

Yarkoni, T., & Westfall, J. (2017). Choosing prediction over explanation in psychology: lessons from machine learning. *Perspectives on Psychological Science: A Journal of the Association for Psychological science*, *12*(6), 1100–1122. https://doi.org/10.1177/1745691617693393

Yi, X. M. (2021). University new students' prognosis of being punished (PBP) for violation and the degree of fear of consequences (DFC) of violation and their influence on the possibility of truancy and cheating. Chinese Journal of Clinical Psychology, 29(3), 572–576. https://doi.org/10.16128/j.cnki.1005–3611.2021.03.026

Ykhlef, H., & Bouchaffra, D. (2017). Induced subgraph game for ensemble selection. *International Journal on Artificial Intelligence Tools*, *26*(1), 1760003. https://doi.org/10.1109/ICTAI.2015.97

Zanette, S., Gao, X. Q., Brunet, M., Bartlett, M. S., & Lee, K. (2016). Automated decoding of facial expressions reveals marked differences in children when telling antisocial versus prosocial lies. *Journal of Experimental Child Psychology*, *150*, 165–179. https://doi.org/10.1016/j.jecp.2016.05.007

Zhang, R. (2019). *Academic dishonesty among primary school students and its correlates*. Hangzhou Normal University. https://kns.cnki.net/KCMS/detail/detail.aspx?dbname=CMFD201902&filename=1019921810.nh

Zhao, L., Mao, H. Y., Compton, B. J., Peng, J. J., Fu, G., Fang, F., Heyman, G. D., & Lee, K. (2022). Academic dishonesty and its relations to peer cheating and culture: a meta–analysis of the perceived peer cheating effect. *Educational Research Review*, *36*, 100455. https://doi.org/10.1016/j.edurev.2022.100455

Zhao, L., Peng, J., Dong, L. D., Li, Y., Mao, H., Compton, B. J., ... & Lee, K. (2022). Effects of test difficulty messaging on academic cheating among middle school children. *Journal of Experimental Child Psychology*, 220, 105417. https://doi.org/10.1016/j.jecp.2022.105417

Zhao, L., Zheng, Y., Mao, H. Y., Compton, B. J., Fu, G., Heyman, G. D., & Lee, K. (2021). Using environmental nudges to reduce academic cheating in young children. *Developmental Science*, *e13108*. https://doi.org/10.1111/desc.13108

The application of artificial intelligence methods in examining elementary school students' academic cheating on homework and its key predictors

ZHAO Li[1]; ZHENG Yi[2]; ZHAO Junbang[3]; ZHANG Rui[1,4]; FANG fang[5]; FU Genyue[1]; KANG Lee[6]

([1] Department of Psychology, Hangzhou Normal University, Hangzhou 311121, China)

([2] Jing Hengyi School of Education, Hangzhou Normal University, Hangzhou 311121, China)

([3] College of Child Development and Education, Zhejiang Normal University, Hangzhou 311231, China)

([4] Hangzhou Xiayan Elementary School, Hangzhou 31112, China)

([5] School of Psychological and Cognitive Sciences and Beijing Key Laboratory of Behavior and Mental Health, Peking University, Beijing 100871, China)

([6] Ontario Institute for Studies in Education, University of Toronto, Ontario M5R 2X2, Canada)

Abstract

Background. Academic cheating has been a challenging problem for educators for centuries. It is well established that students often cheat not only on exams but also on homework. Despites recent changes in educational policy and practice, homework remains one of the most important academic tasks for elementary school students in China. However, most of the existing studies on academic cheating for the last century have focused almost exclusively on college and secondary school students, with few on the crucial elementary school period when academic integrity begins to form and develop. Further, most research has focused on cheating on exams with little on homework cheating. The present research aimed to bridge this significant gap in the literature. We used the advanced artificial intelligence methods to investigate the development of homework cheating in elementary school children and the key contributing factors so as to provide scientific

basis for the development of early intervention methods to promote academic integrity and reduce cheating.

Method. We surveyed elementary school students from Grades 2 to 6 and obtained a valid sample of 2,098. The questionnaire included students' self–reported cheating on homework (the dependent variable). The predictor variables included children's ratings of (1) their perceptions of the severity of consequences for being caught cheating, (2) the extent to which they found cheating to be acceptable, and the extent to which they thought their peers considered cheating to be acceptable, (3) their perceptions of the effectiveness of various strategies adults use to reduce cheating, (4) how frequently they observed their peers engaging in cheating, and (5) several demographic variables. We used ensemble machine learning (an emerging artificial intelligence methodology) to capture the complex relations between cheating on homework and various predictor variables and used the Shapley importance values to identify the most important factors contributing children's decisions to cheat on homework.

Results. Overall, 33% of elementary school students reported having cheated on homework, and the rate of such self–reported cheating behavior increased with grade. The best models with the ensemble machine learning accurately predicted the students' homework cheating with a mean Area Under the Curve (AUC) value of 80.46%. The Shapley importance values showed that all predictors significantly contributed to the high performance of our computational models. However, their importance values varied significantly. Children's cheating was most strongly predicted by their own beliefs about the acceptability of cheatings, how commonly and frequently they had observed their peers engaging in academic cheating, and their achievement level. Other predictors such as children's beliefs about the severity of the possible consequences of cheating

(e.g., being punished by one’s teacher), their beliefs about the effectiveness of cheating deterrence strategies (e.g., working harder) and demographic characteristics, though significantly, were not important predictors of elementary school children’s homework cheating.

Conclusion. This study for the first time examined elementary school students' homework cheating behavior. We used machine learning integration algorithms to systematically investigate the key factors contributing to elementary school students' homework cheating. The results showed that homework cheating already exists in the elementary school period and increases with grade. Advanced machine learning algorithms revealed that elementary school students' homework cheating largely depends on their acceptance of cheating, their peers' homework cheating, and their own academic performance level. The present findings advance our theoretical understanding of the early development of academic integrity and dishonesty and forms the scientific basis for developing early intervention programs to reduce academic cheating. In addition, this study also shows that machine learning, as the core method of artificial intelligence, is an effective method that can be used to analyze developmental data analysis.

***Keywords*：**elementary school students, honesty behavior, academic cheating, cheating on homework, machine learning, prediction, peer behavior